

# Φασματοσκοπία Μαγνητικού Συντονισμού Πυρήνων και Γεωμετρία Αποστάσεων για την πρόβλεψη της τριτοταγούς δομής

Διπλωματική Εργασία Βιοπληροφορικής

Επιβλέπων: Καθηγητής Ι. Εμίρης, Τμήμα Πληροφορικής & Τηλεπικοινωνιών

Συνεπιβλέπων: Δρ Ευαγγελία Χρυσίνα, Εθνικό Ίδρυμα Ερευνών

Η Φασματοσκοπία Μαγνητικού Συντονισμού Πυρήνων (NMR) είναι μια σύγχρονη τεχνική με ραγδαία εξέλιξη και εντυπωσιακά αποτελέσματα, δεδομένου ότι παρέχει τις αποστάσεις μεταξύ ορισμένων πυρήνων ατόμων (Nuclear Overhauser Effect Spectroscopy, NOESY), είτε με ακρίβεια είτε προσεγγιστικά. Το ζητούμενο είναι να κατανοηθεί η τριτοταγής δομή του μορίου, ειδικότερα για εφαρμογές πρόσδεσης ή σύγκρισης δομών.

Μια μαθηματική μοντελοποίηση αποτυπώνει τις αποστάσεις σε έναν πίνακα αποστάσεων και ζητείται να συμπληρωθεί ο πίνακας κατάλληλα [Hav98]. Ισοδύναμα, ορίζουμε βεβαρυμένο γράφο με κόμβους τα άτομα και ακμές που αντιστοιχούν στις αποστάσεις που μετρήθηκαν. Αν βρούμε τρισδιάστατες συντεταγμένες για τους κόμβους του γράφου, τότε ο γράφος εμβυθίστηκε στον Ευκλείδειο χώρο και έχουμε υπολογίσει την τριτοταγή δομή του μορίου [Lau01]. Παρόλο που στην γενική του μορφή πρόκειται για ένα πρόβλημα καθολικής βελτιστοποίησης μιας συνάρτησης με πολλά τοπικά ελάχιστα, υπάρχουν μέθοδοι (και λογισμικά) που το αντιμετωπίζουν σε αρκετές σημαντικές περιπτώσεις και μάλιστα οδήγησαν στο Nobel Χημείας 2002 του Kurt Wüthrich [GMW97, HW].

Η εργασία θα εστιάσει σε συγκεκριμένες κατηγορίες πινάκων αποστάσεων ώστε να υλοποιήσει και να επεκτείνει κατάλληλες αριθμητικές μεθόδους και να τις εφαρμόσει σε μοριακά δεδομένα. Η χρήση τέτοιων μεθόδων σε πίνακες αποστάσεων έχει ήδη δοκιμαστεί επιτυχώς [EN05], αλλά όχι για μεγάλα μόρια.

**Στόχος.** Η αριθμητική βελτιστοποίηση εκτελείται σε ένα αλγεβρικό λογισμικό όπως το Matlab, βασισμένη σε μια σειρά διαταράξεων του πίνακα αποστάσεων, έτσι ώστε να καταλήξουμε σε έναν αποδεκτό πίνακα. Στη μέχρι τώρα δουλειά μας, δεν έχουμε εκμεταλλευτεί την γνώση της ακρίβειας κάθε δεδομένης απόστασης, ούτε το γεγονός πως η κατανομή της πραγματικής απόστασης δεν είναι ομοιόμορφη στα δεδομένα διαστήματα. Τέλος, μια βελτιωμένη μέθοδος διαταραχής μπορεί να χαλαρώσει τον περιορισμό πως η τελική απόσταση οφείλει να βρίσκεται στο εσωτερικό κάθε αντίστοιχου διαστήματος. Παράλληλα, μελετάμε την εμβύθιση γράφων [EM99, ETV09] για την αντιμετώπιση μικρών υποπροβλημάτων με απόλυτη ακρίβεια.

Θα χρησιμοποιηθούν δεδομένα αποστάσεων από πειράματα NMR. Η επιβεβαίωση των θεωρητικών αποτελεσμάτων μπορεί να πραγματοποιηθεί και με την χρήση γνωστών δομών από την Protein Data Bank.

Η παρούσα εργασία θα έχει μια σημαντική ερευνητική συνιστώσα, επομένως όσα περιγράφησαν αποτελούν μια σκιαγράφηση που μπορεί να μεταβληθεί ανάλογα με τα αποτελέσματα της δουλειάς μας. Στοχεύουμε σε μια δημοσίευση σε αναγνωρισμένο διεθνές συνέδριο, η οποία να οδηγήσει τελικά σε δημοσίευση σε διεθνές περιοδικό.

**Γνώσεις.** Απαραίτητη η χρήση ενός αλγεβρικού λογισμικού όπως το Matlab, Maple ή Mathematica. Επιθυμητή η γνώση αλγοριθμικών τεχνικών, κυρίως γραμμικής άλγεβρας ή βελτιστοποίησης.

Οι γνώσεις που θα αποκτηθούν αφορούν στον προγραμματισμό ενός εκ των παραπάνω αλγεβρικών λογισμικών, στη χρήση τους σε συνδυασμό πιθανώς και με άλλα υπάρχοντα λογισμικά βιοπληροφορικής. Επίσης θα αποκτηθούν γνώσεις αλγοριθμικών τεχνικών και πρωτότυπης ερευνητικά εφαρμογής τους σε μοριακά προβλήματα.

## Anafor'es

- [EM99] I.Z. Emiris and B. Mourrain. Computer algebra methods for studying and computing molecular conformations. *Algorithmica, Special Issue on Algorithms for Computational Biology*, 25:372–402, 1999.
- [EN05] I.Z. Emiris and T.G. Nikitopoulos. Molecular conformation search by distance matrix perturbations. *J. Math. Chemistry*, 37(3):233–253, April 2005.
- [ETV09] I.Z. Emiris, E. Tsigaridas, and A. Varvitsiotis. Algebraic methods for counting Euclidean embeddings of rigid graphs. In *Proc. Graph Drawing, Chicago*, LNCS. Springer, 2009. To appear.
- [GMW97] P. Guntert, C. Mumenthaler, and K. Wuthrich. Torsion angle dynamics for NMR structure calculation with the new program Dyana. *J. Mol. Bio.*, 273:283–298, 1997.
- [Hav98] T.F. Havel. Distance geometry: Theory, algorithms, and chemical applications. In P. von Ragué, P. R. Schreiner, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, and H. F. Schaefer III, editors, *Encyclopedia of Computational Chemistry*, pages 723–742. J. Wiley & Sons, 1998.
- [HW] T. Havel and K. Wuthrich. *DISGEO*. program #507, Quantum Chemistry Program Exchange. <http://qcpe.chem.indiana.edu>.
- [Lau01] M. Laurent. *Matrix completion problems*, volume III (Interior - M), pages 221–229. Kluwer, 2001.